47	Ministero della Salute	ITA	2.1 Study design	Page 12 Lines 494-498: It should be stress that case studies especially those referring to acute poisoning episode can be potentially useful for hazard identification and not for risk assessment (suggestion: change the last part of the sentence in 'which makes them potentially relevant for hazard identification'. *EFSA Response:* *The text in line 497 has been amended as following: "…which makes them particularly relevant for hazard identification".* Page 12 Line 499: Clinical trials are not necessarily randomized. If they are, they are called randomized clinical trials. *EFSA Response:* *It has been clarified in the text in line 499 that in the context of this Opinion clinical trials are understood to be randomised.* Page 12 Lines 502-504: The sentence describing observational studies is wrong. The sentence says that in observational studies the exposure is not randomly assigned. The difference described is between randomized and non-randomized clinical studies. In observational studies exposure is not assigned at all! The difference between clinical studies and observational epidemiology is whether the researcher allocates the exposure or not. Observational epidemiology refers to studies where the researcher cannot decide who is exposed and who is not, but can only observe groups with different exposures. *EFSA Response:* *Text in lines 502-504 has been clarified accordingly.* Page 13 Line 531 and following: The inclusion of prevalent cases is a major drawback of (most) cross-sectional studies, particularly for chronic long-term diseases. *EFSA Response:* *In line 535 the following text has been included: "The inclusion of prevalent cases is a major drawback of (most) cross-sectional studies, particularly for chronic long-term diseases.* Page 13 Lines 543-544: Case-control studies require smaller sample size, less time and resources independently from the rareness of the disease. I suggest to change the sentence by adding in line 544 after "compared to prospective studies" the following: and often they are the only viable option prior to "when studying rare outcomes…" *EFSA Response:* *The text has been clarified to include that case-control studies generally require less time and resources and may be considered the only viable option.*

48	Uniformed Services University	USA	2.1 Study design	(Lines 487-516) There is much focus on the idea that intervention studies rank higher than observational studies in terms of design, but bringing the intervention study design into a discussion about pesticides and human health effects is nonsensical as well as irrelevant, since there is no chance ethically to carry out an intervention study, administering pesticides to humans. This type of discussion undermines the message of the authors and demonstrates their lack of understanding of the field of studying the human health effects of pesticide exposure. *EFSA Response:* *It is a fact though that pesticide dossiers are traditionally based on a large number of clinical studies. The description made in this chapter of the opinion is intended to clarify for the reader the differences between intervention studies (in laboratory animals) and observational studies (in humans), considering both the advantages and the disadvantages.*
49	retired epidemiologist	USA	2.1 Study design	2. GENERAL FRAMEWORK OF EPIDEMIOLOGICAL STUDIES ON PESTICIDES 1) The Panel presents intervention studies as an alternative to the "flawed" observations studies. Does the Panel actually believe that intervention studies on chronic outcomes for pesticides is ethical? If so, they should indicate how. *EFSA Response:* *The PPR Panel does not consider that the use of intervention studies on chronic outcomes for pesticides is ethical. In the general introduction chapter, intervention studies have been mentioned mostly to explain the limitations of observational studies. Interventions for potentially harmful chemicals cannot be made and therefore studies that are of less quality by design are used. Within EFSA human observational studies are frequently used for risk assessment (see CONTAM Opinion on Cadmium and lead for example).* 2) The idea of an unknown confounder needs to be explored more fully. First, an unknown confounder must be something that the investigators could not imagine, or could not gather information about in the study. Epidemiologists usually consider as possible confounders all factors that are associated with the disease or outcome of interest and those that are correlated with the exposures of interest. So the "unknown" confounder must be something else. It is, of course, possible for there to be such a factor, but they must be something not suggested in the literature, or are not considered because of the incompetence of the epidemiologists. Is this likely? What is even more disturbing is that there is literally no way to address the "unknown confounder" criticism. No matter how many factors are considered and evaluated as possible confounders, it is still easy to make the charge that there is still some "unknown" confounder out there. To be scientific, the critic should suggest what "confounder" has not been considered. *EFSA Response:* *See also comment #43.* *The PPR Panel considers this true, but there are many examples of studies where lack of control for co-exposures, disease status or social status has created results that have proven difficult to replicate.* 3) The Panel repeatedly gives the impression that the quality of exposure assessment can be determined by the

				study design. This is true only to a minor extent. There are some effects design, but the main impact comes from the actual procedure followed to characterize exposures. *EFSA Response:* *See also comment #44.* *But retrospective exposure assessment is a limitation that is determined by the study design.* 4) Recall bias is a potential in case-control studies, but the Panel gives the impression that it is a given. It is not and the effect can be in either direction. There are many case-control studies where exposure assessment is superior to that in cohort studies. There is a considerable literature on this issue, but the Panel does not review it at all. At a minimum, it would seem the Panel should provide information where case-response bias has occurred for pesticides, i.e., not just the possibility, but documented evidence, as well as studies that have evaluated its occurrence. *EFSA Response:* *See also comment #44.* *In many other branches of epidemiology case-control studies are only taken as suggestive (such as cancer epidemiology). However when results from these studies are in agreement for animal studies, then the combined results provides much stronger evidence than from just relying on the animal studies or the epi studies. Epidemiological studies on pesticides are not a very data rich area, so "documented evidence" to what you suggest most likely does not exist.* 5) • This section should indicate what conditions must occur for non-differential misclassification to bias relative risks away from the null. There are some, but they are not common. Non-differential exposure misclassification cannot bias the relative risk upward for the highest exposure category, as would be necessary to create a biased exposure-response gradient. This is important because the upper category is critical for the slope and anchors the upper end of the risk distribution. In addition to showing how it can occur, the Panel provide documentation of situations where it has. *EFSA Response:* *Blair et al (1993) reported that the reliability and validity of data on pesticide use obtained by recall, often years after the event, have been questioned... Methodologic efforts are needed, however, to assess the actual reliability and accuracy of farmers' reported use of pesticides... The major problems in assessing agricultural pesticide exposure based on information obtained from interviews would result in nondifferential misclassification. This error tends to bias risk estimates toward the null and to dilute exposure response gradient. It may cause false-negative results... Prospective studies of farmers would provide one solution to the misclassification problem, since periodic assessment of exposures would reduce the problem of long-term recall.*
50	German Federal Institute for Risk Assessment (BfR)	DEU	2.2 Population and sample	Please define the term 'risk inflation'. *EFSA Response:*

			size	*Same as effect size inflation (see Annex D).*
51	ECPA	BEL	2.2 Population and sample size	Lines 582-584: "The sample size of a study should be large enough to warrant sufficient statistical power (e.g. 80%). This is the likelihood that an effect of a magnitude that is considered biologically relevant or relevant from a regulatory perspective will also be statistically significant". This point is critical, because it will help prevent regulatory decisions which are based on biological effects that are not adverse or on the path towards an adverse effect. However, it would be useful if the scientific opinion more fully explain what constitutes a biologically relevant effect or provide a reference where EFSA guidance is available. Lines 582-584. The power determination of an epidemiology study is often only done post hoc. A large cohort study may publish analyses of a group of the study subjects, for which a low number may be exposed. The observance and consequence of "effect size magnification" may occur as discussed in Annex D and lines 746 – 779. When evaluating epidemiology data, particular attention needs to focus on which populations are compared to derive meaningful exposure-response associations. As a general observation, farm life is substantially different from life in the general population. Therefore, depending on the focus of the epidemiology investigation, adequate and specific control populations need to be defined, e.g. comparable to the farmer or professional applicator or resident. General population census registers are not necessarily representative for all questions. *EFSA Response:* *The Guidance on biological relevance has been recently published (EFSA Journal 2017; 15(8):4970). Reference to such Guidance has been included in the opinion. The PPR Panel agrees with the other observations you made and which have been already described with a slightly different wording in this chapter.*
52	personal	USA	2.2 Population and sample size	Lines 581-588. The definition of statistical power is actually not correct. Statistical power has nothing to do with biological or regulatory relevance. Power is the probability that you accept the null hypothesis (i.e., no association between pesticide exposure and health outcome) when the alternative hypothesis is true (i.e., an association between pesticide exposure and health outcome). The power of a given study is driven by the sample size, but also the magnitude of the actual effect (assuming one exists) of the exposure and health outcome under study. This may seem like a minor distinction, but this interpretation actually has very large implications for the evaluation and interpretation of null studies, in particular. *EFSA Response:* *Agree. The text has been revised for accuracy.*
53	SYNGENTA	GBR	2.2 Population and sample size	No comment on this description of sampling statistics as applied to epidemiological studies. The reader should also be referred to standard textbooks of epidemiology (e.g. Rothman et al., 2008; Thomas et al., 2009) for a more comprehensive treatment of this subject matter. Thomas, D.C. Statistical Methods in Environmental Epidemiology (2009), Oxford University Press, Oxford, UK.

				EFSA Response: *The references you propose (Rothman et al., 2008 and Thomas et al. 2009) has been included in the Opinion as you suggested (see also comment #61).*
54	Ministero della Salute	ITA	2.2 Population and sample size	Page 14 Lines 589-591: It is not completely true that clinical trials have only one single endpoint, since they have secondary endpoints and mostly they evaluate safety together with efficacy. Better changed it to one or few endpoints and one or few exposures for clinical trials and several endpoints and/or several exposures for observational studies. For example, case-control studies have mostly one endpoint (determined by the type of cases), but generally investigate a large number of exposures *EFSA Response:* *Text in lines 589-591 states that clinical trials are designed and conducted to test one single hypothesis, not one single endpoint.*
55	ECPA	BEL	2.3 Exposure	It would be useful to include a paragraph in this section on the difficulties in assessing historical pesticide exposures. This is particularly important for health outcomes thought to have a long latency period. *EFSA Response:* *The following text has been included at line 1330: "It is particularly challenging to construct an assessment of historical exposures which may deviate from current exposures, in both the range of chemicals and intensity of exposure and also co-exposure to other substances which are not included in the scope of study".*
56	personal	USA	2.3 Exposure	Lines 598-616. The determination of the exact "dose" is one of the real strengths of toxicological studies, and an advantage of these studies over epidemiology. Unfortunately, in the diseases of primary interest listed by the panel, the latency period from exposure to disease is quite long in humans. Therefore, it is unrealistic to think that the exposure assessment in human populations can ever be as precise as it is in animal studies. *EFSA Response:* *Agree. As indicated in this chapter "in the case of pesticides, estimating exposure in a human observational setting is difficult as the dose, its frequency and duration over time and the route of exposure are not controlled and not even well known".*
57	SYNGENTA	GBR	2.3 Exposure	The ability to quantify dose and exposure duration in observational epidemiology studies is typically inadequate in epidemiology studies. Without good exposure data for individuals, the assessment of association with disease becomes uninformative. This topic is treated comprehensively in standard textbooks of environmental epidemiology (e.g. Baker and Nieuwenhuijsen, 2008). Baker, D. and Nieuwenhuijsen, M.J. Environmental Epidemiology: Study methods and applications (2008), Oxford University Press, Oxford, UK. *EFSA Response:*